Nothing is Perfect -

An Insider's Evaluation of Regression Equations Sets for Calculating Flood Flows in Nebraska

Jeffrey T. Shafer, P.E.
Nebraska Department of Natural Resources

Introduction

Regression equations are a quick, effective, and approved (FEMA, 2003) method for computing the 1% chance flow (also referred to as the 100-year flood) for use in Federal Emergency Management Agency flood insurance studies. However, without a thorough knowledge of the regression equation sets strengths and weaknesses, their use may result in wildly varied computed flows along a stream. That type of result will negatively influence the extent of the 1% chance floodplain and cause the public to have doubts about (and potentially disregard) the study results.

Because regression equation use has been an integral part of the success of the **Nebraska Department of Natural Resources Large Area Mapping Initiative**, an evaluation of the Nebraska regression equation sets was necessary. This paper reviews the strengths, weaknesses and applicability of the four primary regression equation sets available for Nebraska streams in order that future flood insurance studies and BFE determinations can made with the most appropriate flow values.

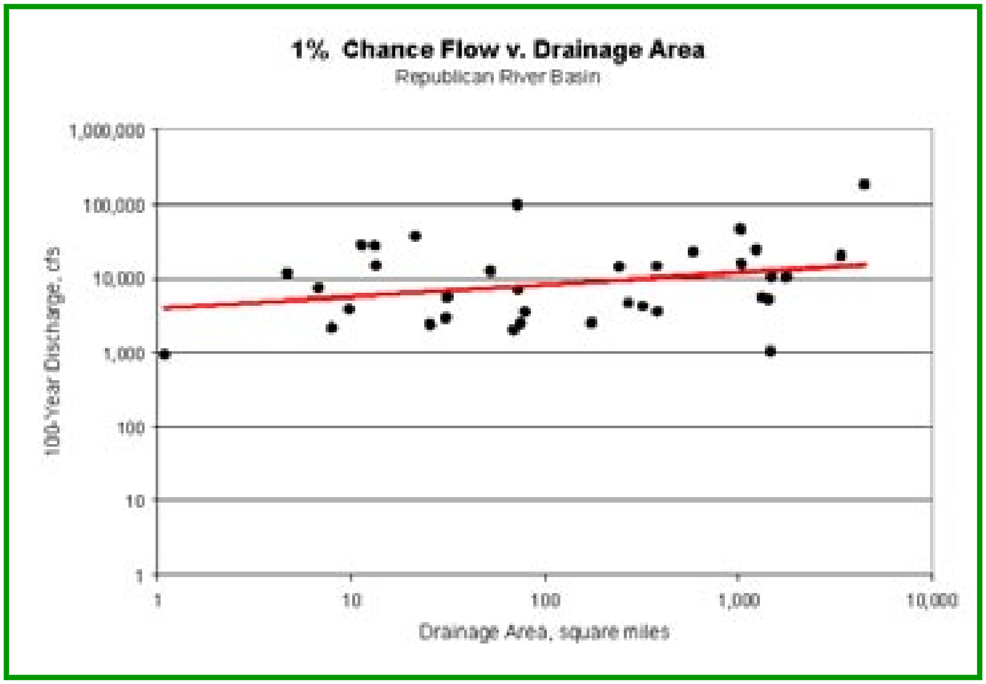
Background

Regression equation sets are a method for estimating the peak flow values for a given watershed. Each equation will relate peak flow and frequency of occurrence to drainage basin characteristics that have the greatest influence on the peak flow. Examples of basin characteristics include drainage area, main channel length, main channel slope, rainfall intensity, mean annual precipitation, and so on. The peak flow and frequency of occurence value is often a Log-Pearson Type III frequency distribution computed using guidelines from Bulletin #17B[1]. The following chart shows a set of 1% chance flows plotted against the basin drainage area. The line running through the data points shows the best fit for that parameter. It is then possible to get a better fit of the line by adding additional basin parameters to the equation.

[1] United State Geological Survey, *Guidelines for Determining Flood Flow Frequency*, Bulletin #17B of the Hydrology Subcommittee, 1982.

Chart 1.

1% Chance Flows v. Drainage Area, Republican River Basin.



Methodology

Each of the **four** current statewide regression equations sets will be given a general evaluation in this paper. In addition a **side-by-side comparison** will be made comparing the 1% annual chance flows (calculated using the equations) along actual stream reaches in different areas of the state.

The general evaluation will use the following criteria:

- ***Are the parameters well defined and easily determined?*** This evaluation will include determining the reproducibility of results, how easily basin parameters used in the equations are determined, if additional tools are necessary beyond a topographic map and pencil, and if there were any problems programming the equation set into ArcView 3.3.
- ***Are the equation result values sensitive to any one parameter?*** This evaluation will look at the changes in results that could occur due to changes in equation parameters. An analysis of percent change in the 1% chance flow will be completed for highly sensitive parameters.
- ***What is the age of the data used in the study?*** It is important to use as much data as possible when calculating flow values using Bulletin 17B techniques. Therefore, this evaluation will look at the cutoff year for data used in the regression formulation.

Regression Equations Sets

Beckman, 1976[2]

This regression equation set divides the state into five regions based on soil types and/or watershed boundaries. Each region has six associated equations for calculation of the 50%, 20%, 10%, 4%, 2%, and 1% annual chance flows. The report was prepared by the U.S. Geologic Survey and sponsored by the Nebraska Department of Roads.

This regression set was easily programmed into ArcView. One issue that did come up during programming was that not all of the region boundaries follow watershed boundaries. This creates the problem of potential significant jumps in calculated discharge along a stream that crosses from one region to another.

The parameters used in the regression set were easily calculated in a geographic information system environment with the exception of contributing drainage area. It is very difficult to accurately calculate the non-contributing portion of a watershed in a geographic information system. This is not unique to GIS calculations: two engineers or technicians, given a set of aerial photographs and topographic maps, may estimate different contributing drainage areas. Therefore, the total drainage area was used for all area calculations.

All of the parameters were well-defined for this equation set. One limitation is that the maps showing the mean minimum January temperature and the normal daily maximum March temperature can only be estimated to the nearest degree.

In region 2, the parameter ($I_{24,50}$ – 3) is raised to a power of 3.731 power, making it very sensitive to minor changes in the estimate of the 24-hour, 50-year intensity. Changing the value from 1.2 to 1.1 would reduce the flow estimate by 28%. Other sensitive values are area, 24-hour, 50-year rainfall intensity, and main channel length in Region 4.

The flow values used in the report were developed using stream gaging record data up through 1972. Bulletin 15[3] methods were used for determining discharges for various return periods from annual the peak flows.

[2] Beckman, E.W. *Magnitude and Frequency of Floods in Nebraska.* USGS Water Resources Investigations 76-109. 1976.
[3] Water Resources Council, *A Uniform Technique for Determining Flood Frequencies; Water Resources Council Bulletin 15*, 1967.

Cordes and Hotchkiss, 1993[4]

This report was developed at the University of Nebraska Civil Engineering Department and sponsored by the Nebraska Department of Roads. One of the purposes of the study was to update the regional regression equations of the State of Nebraska developed by Beckman. As with the Beckman equations, this regression equation set divides the state into five regions based on soil types and/or watershed boundaries. Each region has six associated equations for calculation of the 50%, 10%, 2%, 1%, 0.5%, and 0.2% annual peak flows.

This regression set was easily programmed into ArcView. Since the region boundaries are the same as those developed by Beckman, the Cordes equations have the same issue of jumps in calculated discharge across region boundaries.

The parameters used in the regression set were easily calculated in a geographic information system environment with the exception of contributing drainage area. As with the Beckman equation, it is difficult to estimate the portion of a watershed that contributes to drainage area.

All of the parameters were well defined for this equation set. Limitations in the maps for certain parameters also occur for this equation set. The mean annual precipitation map has 4-inch intervals making it difficult to estimate the value to the tenth of an inch and the map showing the normal daily maximum March temperature only allows estimates to the nearest degree. One parameter, mean normal daily maximum March temperature, isn't intuitively linked to peak flows in Nebraska.

In region 5, the parameter $I_{24,2}$ is raised to a power of 10.491, making it exceptionally sensitive to minor changes in the estimate of the 24-hour, 2-year rainfall intensity. Changing the value from 2.2 to 2.1 would reduce the flow estimate by 40%. In region 3, the parameter P is raised to a power of 5.581, making it very sensitive to minor changes in the estimate of the mean annual precipitation. Changing the value from 17 to 16 would reduce the flow estimate by 29%. In region 2, the parameter SN10 is raised to a power of 2.615, making it sensitive to minor changes in the estimate of the 10% probability equivalent snow moisture content as of March 15. Changing the value from 1.5 to 1.4 would reduce the flow estimate by 16%. Other sensitive values are main channel slope in Region 2, and normal daily March temperature and main channel length in Region 4.

The peak flow values used in the report were developed using stream gaging data up through 1991, 19 more years of record than was used to develop the Beckman equations. Bulletin 17B methods were used for determining the flows.

[4] Cordes, K.E. and R.H. Hotchkiss. *Design Discharge of Culverts.* NDOR Research Project No. RES-1. 1993.

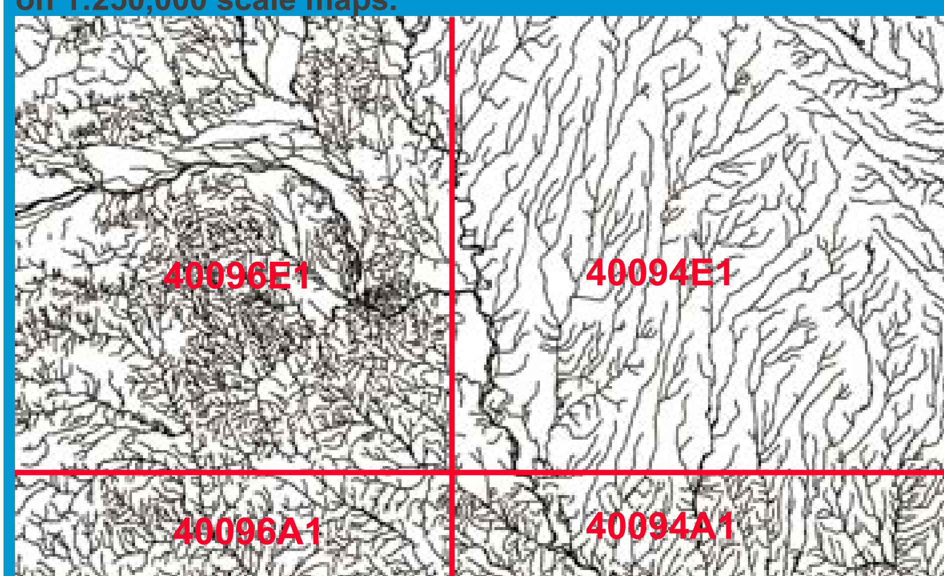
Soenksen et al, 1999[5]

This report was developed by the U.S. Geological Survey and sponsored by the Nebraska Department of Roads. The purpose of the study was to update peak-flow frequency analyses for selected streamflow-gaging stations and to develop a new set of peak-flow frequency relations for ungaged streams. This regression equation set divides the state into seven regions based on soil permeability and watershed boundaries. Each region has eight associated equations for calculation of the 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2% annual chance flows.

This regression equation set was not easily programmed into ArcView because of some of the parameters used in the equations. The parameters used in the equations were intended to be calculated by BasinSoft, a GIS-based program based on Arc/Info. BasinSoft was not readily accessible for this study and therefore interpretations had to be made about variable inputs. In addition, there is a large scale dependency of some parameters. Examples include:

- As with Beckman and Cordes and Hotchkiss, it is very difficult to accurately calculate the non-contributing portion of a watershed in a geographic information system. Therefore the total drainage area was used for all area calculations.
- 1:250,000-scale hydrography maps were used for calculating the drainage frequency and stream density. In Nebraska, the densities of digitized streams vary greatly from one 1:250,000-scale map to the next. **(See Figure 1)**
- The basin slope value is almost impossible to duplicate because of the vague instructions included in the report for calculating it.

Figure 1.
Comparison of Stream Density for two Nebraska counties on 1:250,000 scale maps.



The region boundaries are not clearly defined for the Soenksen equations. There are places where the Northeastern Region overlaps the Eastern Region and where the Northeastern Region overlaps the Central and South-Central Region. In locations within overlapping regions, the two applicable equations may result in calculated discharges that differ by an order of magnitude. In areas near the high permeability region, the user must check the average permeability of the 60-inch soil profile to determine if

[5] Soenksen, P.J. et al. *Peak-Flow Frequency Relations and Evaluation of the Peak-Flow Gaging Network in Nebraska.* USGS Water Resources Investigations Report 99-4032. 1999.

you are in fact in that region or are in one of the surrounding regions. One important note is that without a GIS to calculate some of the parameters, it would be virtually impossible to calculate flows.

In the Central and South-Central Region, the parameter (TTP-2) is raised to a power of 3.83, making it exceptionally sensitive to minor changes in the estimate of the 2-year, 24-hour intensity. Changing the parameter value from 0.5 to 0.4 would reduce the flow estimate by 58%. Other sensitive values are soil available water content in the high permeability region and main channel slope in the Upper Republican Region.

There are also issues with this set of equations in the Blue River Region when comparing the flows for all return periods. For example, the following table **(Table 1)** of flows (for a watershed with an area of 94 square miles) shows that how the parameters are interpreted can greatly affect the perception of their validity:

Table 1.
Flows Calculated with Soenksen Equation, for a Watershed in the Big Blue River Region

Percent Chance Flow	Calculated Flow
50%	12,561 cfs
20%	12,362 cfs
10%	13,293 cfs
4%	7,803 cfs
2%	9,971 cfs
1%	12,278 cfs
0.5%	14,859 cfs
0.2%	18,595 cfs

The peak flow values used in the report were developed using data up through 1993, 2 more years of record than Cordes and Hotchkiss. Bulletin 17B methods were used for determining the peak flows.

Strahm and Admiraal, 2004[6]

This report was developed at the University of Nebraska Civil Engineering Department and sponsored by the Nebraska Department of Roads. The purpose of the study was to modify the Soenksen regional regression equations so that flows could be programmed into a GIS. As with Soenksen, the state is divided into 7 regions based on the soil permeability or watershed boundaries. With a few exceptions, the region boundaries line up with Soenksen. Each region has eight associated equations for calculation of the 50%, 20%, 10%, 4%, 2%, 1%, 0.5%, and 0.2% annual chance flows and eight more for doing the same calculations on watersheds less than 10 square miles in area.

This regression set was easily programmed into ArcView. The parameter definitions were clearly described in the report. The parameters used in the regression set were easily calculated in a geographic information system environment with the exception of contributing drainage area. As with the other equations, it is difficult to estimate the portion of a watershed that contributes to drainage area.

The regions for this equation were defined similarly to the regions defined for the Soenksen equations, although there are a few areas that fall into a different region. Therefore the Strahm and Admiraal equations, as with the Soenksen equations, suffer from ambiguity with respect to overlapping regions and the poorly-defined high permeability region. As with the Soenksen equations, without a GIS to do some of the parameter calculations, it would be virtually impossible to calculate peak flows.

[6] Strahm, B.J. and D.M. Admiraal. *Regression Equations.* NDOR Research Project No. SPR-1(2) P541. 2004.

The calculated flows are sensitive when the drainage area is less than 10 square miles. The results will vary widely between the less than 10 square mile equation and the complete basin equation.

No peak flow values were developed using stream gage records in this report. The peak flow values used in the report were taken from Soenksen.

Results

Side-by-Side Comparison

The 1%-chance flow was calculated at intervals of approximately 1 mile along a stream and the results plotted as separate graphs for five basins.

Chart 2 depicts discharges calculated for Blackwood Creek, which is primarily in Hayes County. The discharge calculated using the Beckman equation correlates almost linearly with drainage area, with a major exception where a portion of the streamline falls within an adjacent zone. Discharges calculated from the other three equations increase rapidly with increasing drainage area, then decrease for drainage areas between 10 and 50 square miles, then increase slowly. One exception is the Cordes and Hotchkiss discharges, which shows a rapid increase at 50 square miles where the drainage basin crosses a zone boundary.

Chart 2.
Blackwood Creek in Hayes County.

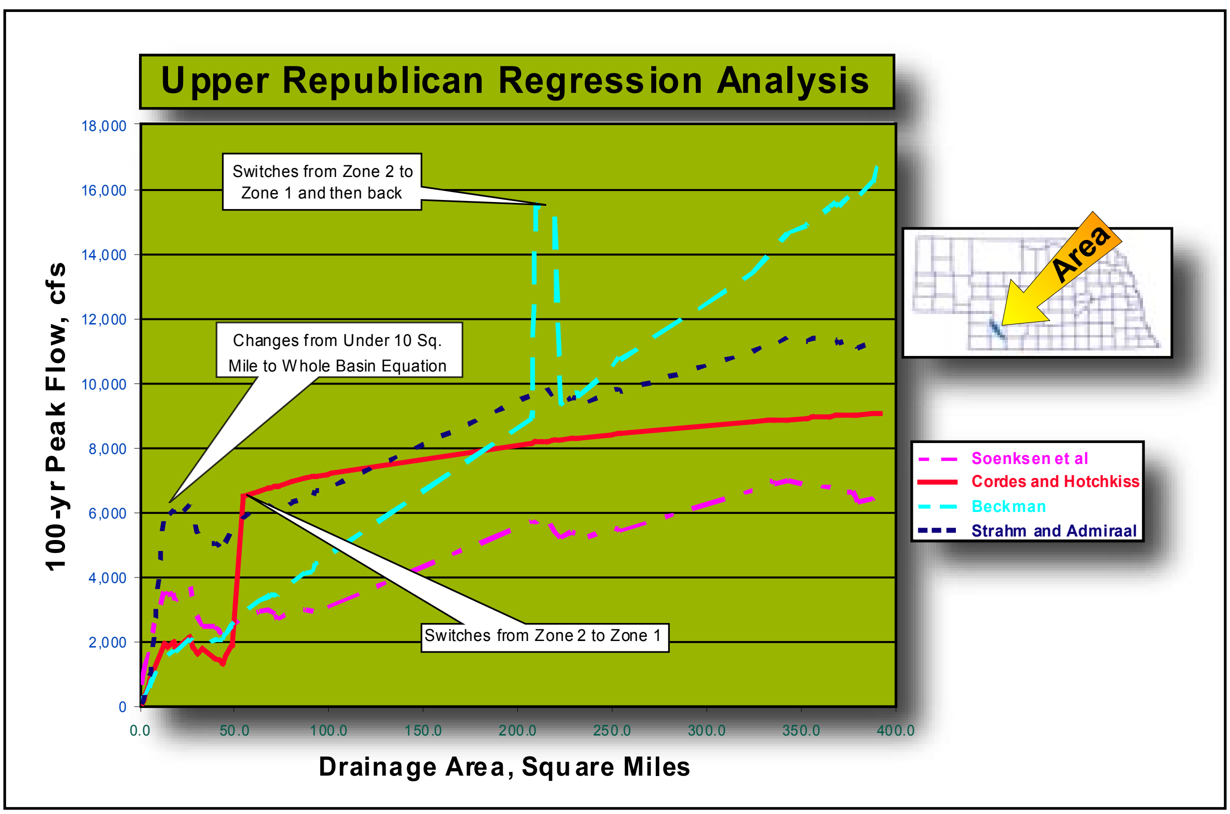


Chart 3.
South Omaha Creek in Thurston County.

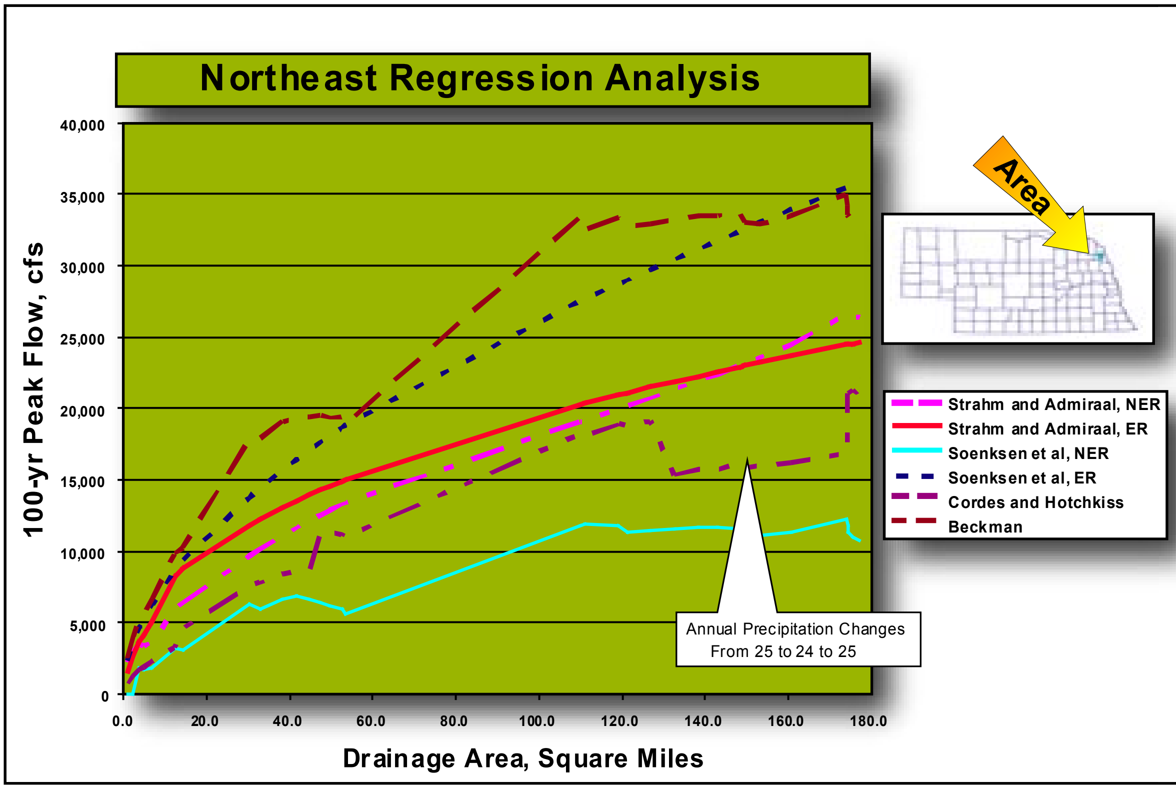


Chart 3 shows results for South Omaha Creek in Thurston County. The creek is located where the Eastern Region and the Northeastern Region overlap for the Soenksen and Strahm equations. Therefore a total of six discharges have been graphed for this creek. The calculated discharge behaves predictably for all six equations, with the exception of the Cordes equation. The Cordes equation decreases greatly where the annual precipitation changes between 25 to 24 inches per year.

Chart 4.
Big Indian Creek in Gage County.

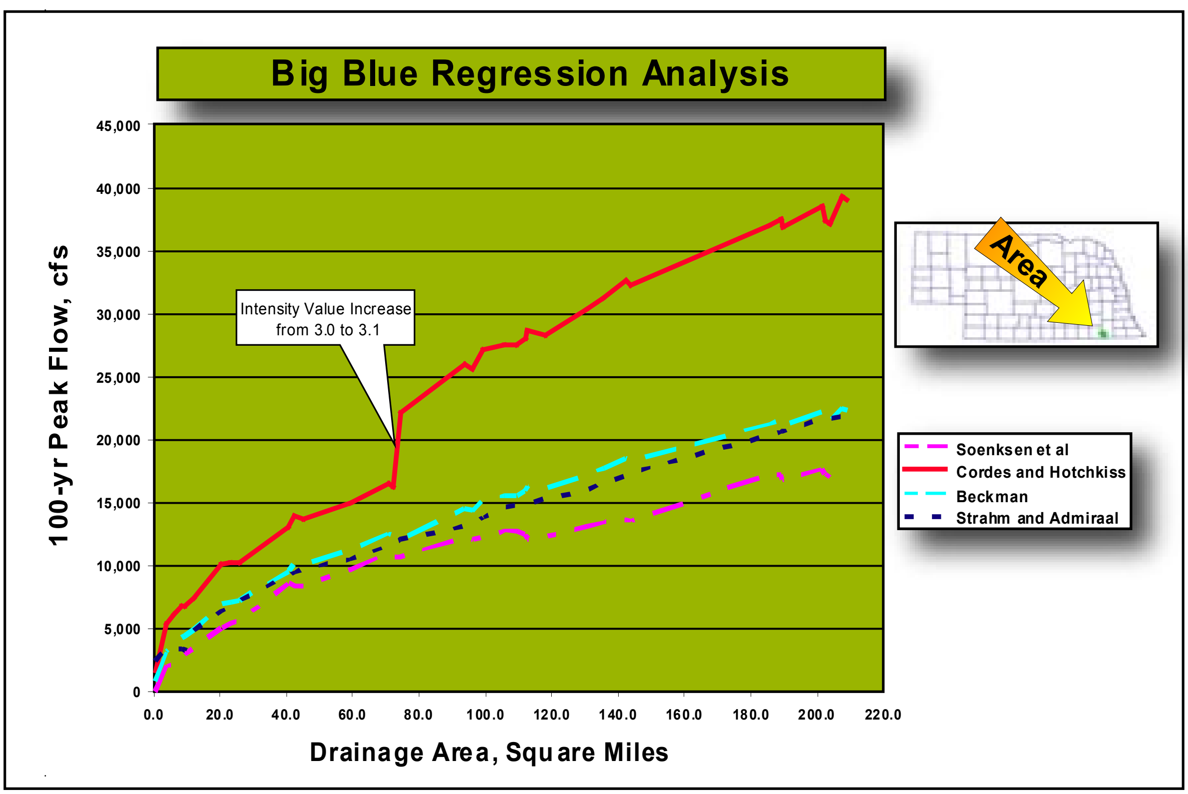


Chart 4 compares the four regression equations for Big Indian Creek in Gage County. The discharge calculated using the Cordes equation is somewhat larger than the discharge calculated from the other three equations. Of special note is the large increase in discharge calculated where the two-year 24-hour rainfall intensity increases from 3.0 to 3.1 inches. This is a startling example of the sensitivity of this particular parameter.

Chart 5.
Spring Creek in Greeley County.

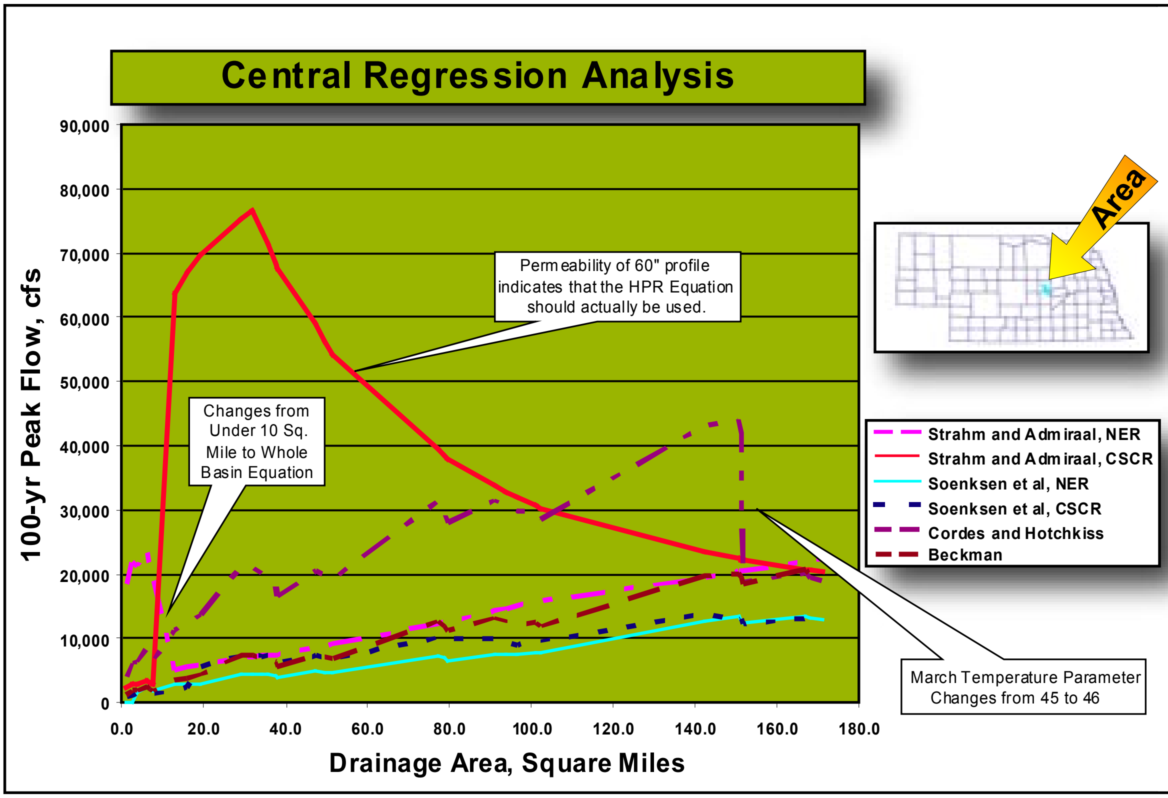


Chart 5 documents the results for Spring Creek in Greeley County. The creek is located where the Northeastern Region and the Central and South-Central Region overlap for the Soenksen and Strahm equations. The first thing to note is that the shape of the curve for Strahm Central and South-Central equation results are very odd. This is likely because the soil permeability in the upper part of the reach is greater that 4 inches/hour and therefore the High Permeability Region equation should be used. The other jumps along the results can be attributed to parameter sensitivity and changing from the under 10 square mile equations to the basin-wide equation set.

Chart 6.
Lone Tree Creek in Dawes County.

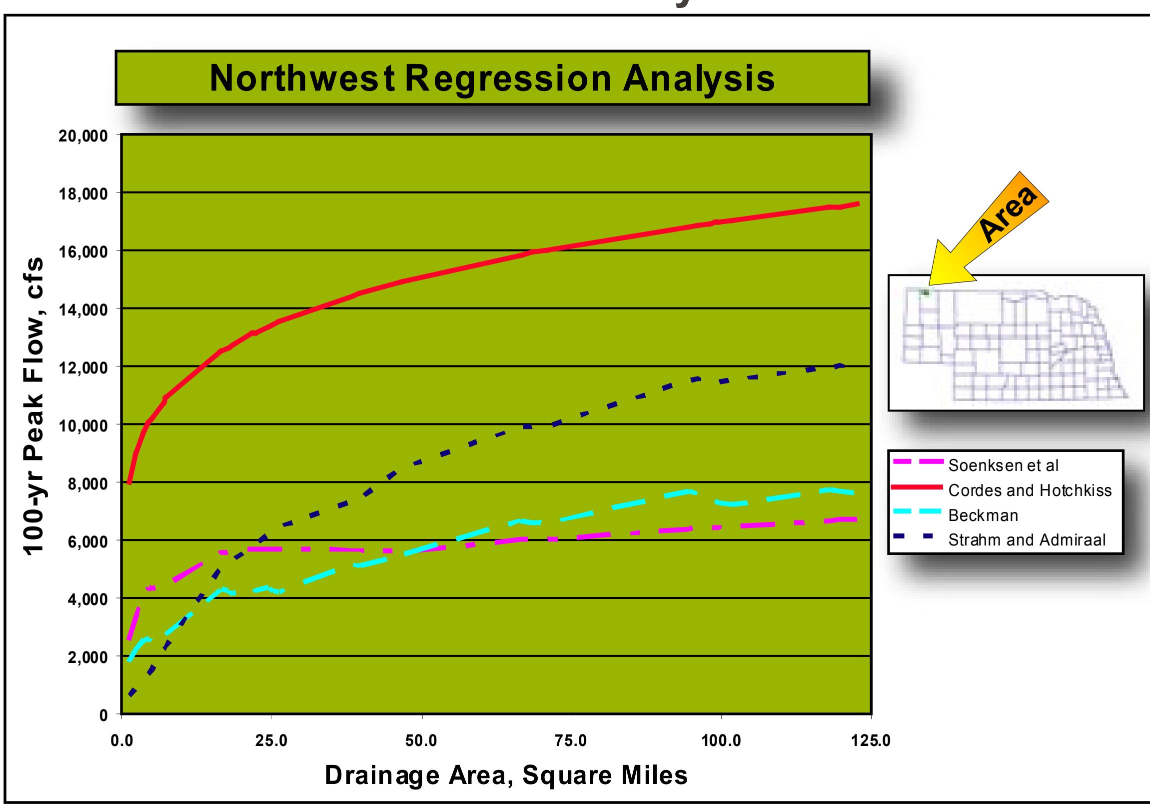
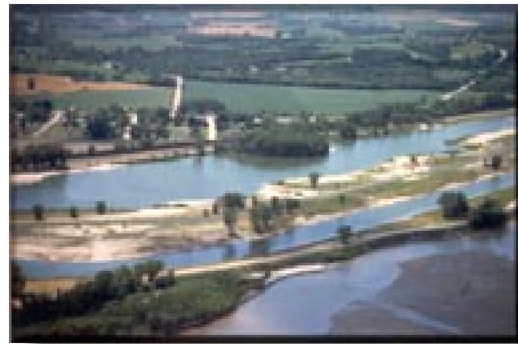


Chart 6 shows results for Lone Tree Creek in Dawes County. All four equations behave well along this creek. Unfortunately, the magnitude of the calculated flows varies widely from equation to equation. The highest calculated discharges (Cordes equation) are approximately two and a half times the lowest calculated discharges.

Summary



All of the regression sets available for use in Nebraska have deficiencies in various regions and should only be used with an understanding of what those deficiencies are. The Strahm and Admiraal equation set may be the best available for many situtations, but it is not without limitations. All future Nebraska Department of Natural Resources flood insurance studies will use a regression equation set that is determined on a study by study basis. The strengths and weaknesses of each set, the results along multiple streams in the study area, and past experience of the modeling team will all be used to determine the regression set or sets for the study.

Appendix A

Beckman

1% Chance Peak Flow Regression Equations

Region 1

$$Q_{100} = 996.78 A_C^{0.624} (P-13)^{0.588} L^{-0.512}$$

Region 2

$$Q_{100} = 0.55 A_C^{0.872} S^{1.063} (I_{24,50} - 3)^{3.731}$$

Region 3

$$Q_{100} = 1162 A_C^{0.843} (T_3 - 37)^{3.636} L^{-0.671}$$

Region 4

$$Q_{100} = 31454 A^{1.724} (I_{24,50} - 5)^{1.365} L^{-2.184}$$

Region 5

$$Q_{100} = 63.87 A_C^{0.680} (T_1 - 11)^{0.741} S^{0.572}$$

Where

Q_{100} = Annual 1% chance peak flow

A_c = Contributing drainage area, square miles

A = Total drainage area, square miles

P = Mean annual precipitation, inches

L = Main stream length, miles

S = Main stream slope, feet/mile

$I_{24,50}$ = Maximum 24-hour, 50-year rainfall, inches

T_3 = Normal daily maximum March temperature, degrees

T_1 = Mean minimum January temperature, degrees

Cordes and Hotchkiss

1% Chance Peak Flow Regression Equations

Region 1

$$Q_{100} = 23553 A_C^{0.170} (P - 13)^{-1.011}$$

Region 2

$$Q_{100} = 0.0816 A_C^{1.051} S^{2.119} SN10^{2.415}$$

Region 3

$$Q_{100} = 0.00000326 A_C^{0.681} S^{0.497} P^{5.581}$$

Region 4

$$Q_{100} = 31008 A^{1.433} (T_3 - 43)^{-1.876} L^{-1.648}$$

Region 5

$$Q_{100} = 0.00335 A_C^{0.615} S^{0.621} I_{24,2}^{10.491}$$

Where

- Q_{100} = Annual 1% chance peak flow
- A_c = Contributing drainage area, square miles
- P = Mean annual precipitation, inches
- L = Main stream length, miles
- S = Main stream slope, feet/mile
- $I_{24,2}$ = Maximum 24-hour, 2-year rainfall, inches
- T_3 = Normal daily maximum March temperature, degrees
- SN10 = 10% probability equivalent snow moisture content as of March 15

Soenksen et al

1% Chance Peak Flow Equations

High Permeability Region

Standard Analysis

$$Q_{100} = 119CDA^{0.777}(MAP-15)^{0.787}AWC^{1.56}MCS^{0.860}$$

Composite Analysis

$$Q_{100} = 776CDA^{0.828}(MAP-15)^{0.741}AWC^{2.07}DF^{0.641}MCS^{0.941}$$

Northern and Western Region

$$Q_{100} = 35.2CDA^{0.213}BS^{0.589}(MAP-12)^{0.643}$$

Northeastern Region

$$Q_{100} = 3000TDA^{0.583}SF^{-0.573}DF^{0.384}PLP^{-0.223}$$

Central and South-Central Region

$$Q_{100} = 104TDA^{0.914}SF^{-0.560}RR^{1.93}(TTP-2)^{3.83}$$

Eastern Region

$$Q_{100} = 242CDA^{0.485}BS^{0.343}PLP^{-0.474}$$

Upper Republican River Region

$$Q_{100} = 9.45CDA^{0.613}MCS^{1.54}CR^{-0.905}$$

Big Blue River Region

$$Q_{100} = 764TDA^{0.483}MSS^{0.656}SF^{-0.382}SD^{0.601}$$

Where

- Q_{100} = Annual 1% chance peak flow
- CDA = Contributing Drainage Area, square miles
- MAP = Mean Annual Precipitation, inches
- AWC = available Water Capacity of 60-inch soil profile, inches per inch
- DF = Drainage Frequency, in first order streams per mile
- MCS = Main Channel Slope, ft/mile
- BS = Basin Slope, ft/mile
- TDA = Total Drainage Area, square miles
- PLP = permeability of the least permeable layer, inches/hour
- SF = Shape Factor, dimensionless
- RR = Relative Relief, feet per mile
- TTP = 2-year, 24-hour precipitation, inches
- CR = Compactness Ratio, dimensionless
- MSS = Average Maximum Soil Slope, percent
- SD = Stream Density, miles per square mile

Strahm and Admiraal
1% Chance Peak Flow Equations

Big Blue Region

Less than 10 square mile basins

$$Q_{100} = 154CDA^{0.186}PLP^{-0.665}MCS^{0.473}$$

Complete Basin Analysis

$$Q_{100} = 444CDA^{0.436}SF^{0.024}MSS^{0.836}$$

Eastern Region

Less than 10 square mile basins

$$Q_{100} = 365CDA^{0.563}BS^{0.287}P60^{-0.536}$$

Complete Basin Analysis

$$Q_{100} = 905CDA^{0.381}BS^{0.163}PLP^{-0.491}$$

Northeastern Region

Less than 10 square mile basins

$$Q_{100} = 8335CDA^{0.285}SF^{-0.556}PLP^{0.728}$$

Complete Basin Analysis

$$Q_{100} = 16CDA^{0.497}BS^{0.629}MSS^{0.368}$$

Central and South-Central Region

Less than 10 square mile basins

$$Q_{100} = 931CDA^{0.290}BS^{0.245}SF^{-0.430}$$

Complete Basin Analysis

$$Q_{100} = 870CDA^{0.209}MCS^{0.107}P60^{-0.187}$$

Upper Republican Region

Less than 10 square mile basins

$$Q_{100} = 4.6CDA^{1.001}MCS^{1.460}P60^{-0.769}$$

Complete Basin Analysis

$$Q_{100} = 91CDA^{0.436}MCS^{0.986}CR^{-0.187}$$

Northern and Western Region

Less than 10 square mile basins

$$Q_{100} = 1722CDA^{0.815}BS^{-0.553}MSS^{0.663}$$

Complete Basin Analysis

$$Q_{100} = 7512CDA^{0.131}RR^{-0.459}PLP^{-0.182}$$

High-Permeability Region

$$Q_{100} = 19CDA^{0.826}BS^{0.181}PLP^{0.297}$$

Where

- Q_{100} = Annual 1% chance peak flow
- CDA = Contributing Drainage Area, square miles
- MSS = Average Maximum Soil Slope, percent
- PLP = permeability of the least permeable layer, inches/hour
- MCS = Main Channel Slope, ft/mile
- SF = Shape Factor, dimensionless
- BS = Basin Slope, ft/mile
- P60 = Permeability of 60-inch profile, in/hr
- CR = Compactness Ratio, dimensionless
- RR = Relative Relief, feet per mile